Atalanta (Mai 1992) 23(1/2):113-118, Würzburg, ISSN 0171-0079

Molekulare Systematik – Verschiedene Techniken und ihre Aussagekraft

von

RALF SOMMER

eingegangen am 12.XI.1991

Einleitung

Die Aufgabe der systematischen Zoologie liegt in der Klassifikation der Organismen aufgrund phylogenetischer Zusammenhänge. Neben morphologischen sind vor allem im letzten Jahrzehnt auch molekularbiologische Methoden herangezogen worden, um verwandtschaftliche Zusammenhänge zwischen verschiedenen Arten aufzuzeigen. Die sogenannte molekulare Systematik hat dabei in kurzer Zeit mit verschiedenen methodischen Ansätzen eine große Datenfülle geliefert. Die dabei zum Einsatz kommenden Methoden sind sehr unterschiedlich und somit ist ihre Aussagekraft nicht äquivalent. Deshalb sollen im folgenden einige wichtige Methoden vorgestellt und ihre Aussagekraft verglichen werden.

Nukleinsäuren und Proteine als Bausteine der Zellen

Die Zelle als Einheit aller Lebewesen ist aus vielen verschiedenen chemischen Substanzen aufgebaut. Alle wichtigen Bestandteile lassen sich in vier Stoffklassen einteilen: Nukleinsäuren, Proteine, Fettsäuren und Kohlenhydrate (siehe ALBERTS & JAENIKE, 1986, für Übersichtsartikel).
Alle vier Stoffklassen werden in der Molekularbiologie untersucht, für taxonomische Arbeiten sind aber nur Nukleinsäuren und Eiweiße (Proteine) relevant. Eiweiße sind die Funktionsträger in lebenden Zellen, sie führen Reaktionen im Stoffwechsel durch oder haben Transportfunktion. Darüberhinaus sind sie für die Beweglichkeit und vieles mehr in den Zellen verantwortlich. Die Nukleinsäuren dagegen sind Informationsträger. Sie enthalten die Information zur Bildung der Eiweiße und anderer biologisch wichtiger Moleküle. Darüberhinaus sind sie auch für die Vererbung der Eigenschaften auf die Nachkommen eines Individuums verantwortlich. Aus chemischer Sicht sind Nukleinsäuren und Proteine sehr verschieden. Gemeinsam ist ihnen aber, daß sie aus langkettigen Verknüpfungen von Einzelbausteinen aufgebaut werden.

Betrachten wir die Proteine: Es gibt 20 verschiedene Bausteine, mit anderen Worten, Proteine besitzen ein Alphabet aus 20 Buchstaben. Diese Buchstaben, die sogenannten Aminosäuren, sind universell, d.h. sie sind identisch bei Menschen und Bakterien.
Jedes Eiweiß besteht aus einer Verknüpfung von vielen, meist über 100 einzelnen Aminosäuren. Die Reihenfolge der Aminosäuren sind für alle Proteine spezifisch. Auch ein Protein, welches in zwei verschiedenen Arten die gleiche Funktion ausübt, z.B. den Abbau von Alkohol, hat eine unterschiedliche Reihenfolge der Aminosäuren in den verschiedenen Arten, wenn auch, je nach Verwandtschaft, mehr oder weniger große Übereinstimmungen

bestehen. Aufgrund der enormen Länge der Eiweiße und der Zahl von 20 verschiedenen Bausteinen gibt es niemals zwei absolut identische Proteine in verschiedenen Arten.

Kommen wir zu den Nukleinsäuren, genauer zur Desoxyribonukleinsäure (DNS): Auch sie ist chemisch gesehen eine langkettige Verbindung. Ihr Alphabet hat aber nur vier Buchstaben, die sogenannten Mononukleotide. Die Reihenfolge der Mononukleotide in der DNS bestimmt die Reihenfolge der Aminosäuren in den Proteinen. Man bezeichnet deshalb den Abschnitt der DNS, der die gesamte Information für die Bildung eines Proteins enthält als "Gen" Kennt man die "Sequenz", d.h. die Reihenfolge der Mononukleotide eines Gens, so kann man daraus die Sequenz der Aminosäuren im zugehörigen Protein ableiten (siehe Abb. 1).

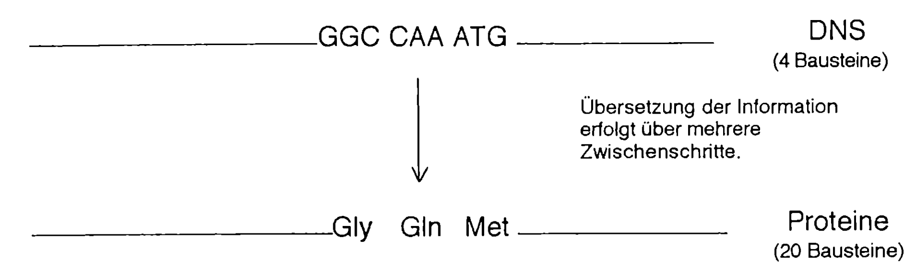


Abb.: 1: Vereinfachtes Schema des molekularbiologischen Dogmas. Die obere Linie stellt die DNS dar, die aus nur 4 Bausteinen, Adenosin- Guanosin- Cytidin- und Thymidintriphosphat besteht. In einem komplizierten Prozeß über mehrere Zwischenschritte werden die Gene in Proteine übersetzt. Die Proteine enthalten 20 verschiedene Bausteine, die in jedem Protein in einer charakteristischen Reihenfolge auftreten.

In der molekularen Systematik untersucht man heute einzelne Eiweiße oder zugehörige Gene. Den größten Informationsgehalt über ein Eiweiß oder ein entsprechendes Gen erhält man, wenn man die Reihenfolge seiner Bausteine bestimmt. Diese optimale Situation ermöglicht einen Vergleich der Proteine verschiedener Arten oder Populationen. Man kann Punkt für Punkt jede einzelne Position zwischen verschiedenen Arten vergleichen und die Anzahl der Übereinstimmungen als Maß für die Verwandtschaft heranziehen.

Nun aber zu den Techniken: Zuerst müssen Proteine oder DNS aus dem Gewebe der Organismen isoliert werden. Dies gelingt recht gut, aber man erhält ein Gemisch von vielen verschiedenen Proteinen, obwohl man nur an einem oder wenigen Proteinen interessiert ist. Für die DNS sieht es ähnlich aus: Man isoliert die Gesamtheit der DNS, will aber nur ein Gen analysieren, und dies macht oft nur einen millionstel Anteil der Gesamt-DNS aus. Will man die Reihenfolge der Bausteine der Proteine oder der DNS bestimmen (man nennt diesen Vorgang "Sequenzierung"), so muß man das Protein bzw. das Gen reinigen und von allen ungewollten Anteilen antrennen. Und an dieser Stelle setzt das Problem ein: Die Reinigung von Proteinen ist ein sehr schwieriges und langwieriges Unterfangen, daß man nicht routinemäßig für viele Proben durchführen kann. Es ist aber auch heutzutage gar nicht mehr nötig, da die Isolierung eines Gens aus der Gesamt-DNS im Zeitalter der Gen-

technologie sehr schnell und einfach möglich ist. In vielen Fällen kann man das gewünschte Gen schon "über Nacht" isolieren. Da, wie oben bereits beschrieben, aus der Gensequenz die Proteinsequenz ableitbar ist, arbeitet man vorzugsweise mit der leicht handhabbaren DNS.
Unabhängig von Sequenzuntersuchungen kann man auch mit anderen Methoden versuchen, die Eigenschaften von Molekülen zu erforschen. Es gibt eine Vielzahl von Methoden, mit denen man Informationen über ein gewünschtes Gen erhalten kann, ohne es gereinigt zu haben. Diese meist elektrophoretischen Methoden werden vor allem bei den Proteinen seit längerer Zeit angewendet. Allerdings sind die dabei erzielten Aussagen nicht annähernd so informationsreich wie die heute möglichen Sequenzinformationen. Evolutionäre Betrachtungen müssen also immer mit Rücksicht auf die benutzten Methoden beurteilt werden. Im folgenden sollen einige Methoden vorgestellt und ihr Informationsgehalt beschrieben werden.

Proteinanalysen

Isoliert man Eiweiße aus einem Gewebe, so erhält man ein Gemisch aus vielen einzelnen Proteinen. Diese werden im nächsten Schritt voneinander getrennt. Bei einer Auftrennung nutzt man die unterschiedlichen Eigenschaften der Proteine: Alle Proteine haben ein spezifisches Gewicht und eine spezifische Ladung. Dies resultiert aus der unterschiedlichen Aminosäurezusammensetzung der Proteine. Die Ladungs- und Gewichtseigenschaften der Proteine können nun benutzt werden, um die Eiweiße eines Extraktes voneinander zu trennen. In sogenannten Stärkegelen sie bestehen aus hydrolisierter Kartoffelstärke – wandern die Proteine in einem elektrischen Feld. Sie wandern je nach ihren Ladungs- und Gewichtseigenschaften unterschiedlich weit. Nach dieser Auftrennung kann man ganz bestimmte Proteine im Gel mit Hilfe von spezifischen Farbreaktionen nachweisen. Dabei nutzt man die natürlichen Stoffwechselreaktionen der Proteine und vollzieht die entsprechenden Reaktionen im Gel nach.
Der Informationsgehalt derartiger elektrophoretischer Experimente beschränkt sich auf das Laufverhalten des Eiweißes im Gelsystem. Man erhält also genau eine Informationseinheit, die man zum Vergleich zwischen verschiedenen Arten heranzieht. Welchen Problemen und Fehlinterpretationen man unterliegen kann, wenn man nur das Gesamtlaufverhalten als Information analysiert, sei an einem Beispiel gezeigt:

	1	2	3	4	5	6	7	8	9	10	Summe
A	+	n	n		+	n		n	+	n	1+
B	+	+	n	+	+	n		n	+	n	4+
C		n	+	n	+	n	n		+	n	1+

Ausstausche	A - B	2
	A - C	5
	B - C	6

Abb. 2: Das Beispiel zeigt Proteine, für die nur 10 Bausteine betrachtet werden. Zur weiteren Vereinfachung sei nur der Faktor Ladung mit +, – und n (für neutral) dargestellt.

Zwischen A und B gibt es nur zwei Austausche, aber eine stark unterschiedliche Gesamtladung. A und C besitzen die gleiche Ladung, aber sechs Unterschiede in der Sequenz. In einem Stärkegel würden die Proben A und C gleichlaufen, und B aber stark unterschiedlich. Betrachten wir die Zahl der Austausche zwischen zwei Formen als Verwandtschaftsgrad, so wird deutlich, daß das identische Laufverhalten zwischen A und C nicht mit den verwandtschaftlichen Beziehungen übereinstimmt. Austausche an verschiedenen Positionen können sich kompensieren, so daß evolutionäre Unterschiede verborgen bleiben; oder aber bei wenigen gleichförmigen Veränderungen so addieren, daß größere evolutionäre Distanzen vorgetäuscht werden.

Es existieren noch zahlreiche andere Trennungsverfahren, mit denen die Proteine nach Kriterien getrennt werden. Bei all diesen Methoden liegt aber das Limit im geringen Informationsgehalt, den man über das gewünschte Protein erhält. Dennoch hat man lange Zeit, auch mangels von Alternativen, mit den oben beschriebenen Methoden gearbeitet. Dabei sind z.T. richtige Ergebnisse erzielt worden, die auch moderneren Techniken standhalten. Oft aber trifft der in Abb. 2 beschriebene Fall zu, und die ermittelten evolutionären Distanzen sind mit genaureren Methoden nicht nachvollziehbar (siehe dazu z.B. die Einleitung bei HILLIS & MORITZ, 1990).

DNS-Analysen

In der Gentechnologie ist es heute möglich, mit vertretbarem Aufwand die Sequenz der Bausteine eines Gens zu bestimmen. Damit hat man die Möglichkeit, Position für Position zwischen verschiedenen Arten zu vergleichen; man erhält also eine Vielzahl von nützlichen Einzelinformationen pro Gen. Die Methode selbst braucht hier nicht vorgestellt zu werden, da es sich nur um molekularbiologische Details handelt. Wichtig ist nur das Ergebnis: Man erhält die Reihenfolge der Mononukleotide der DNS und kann damit die Reihenfolge der Aminosäuren des Proteins ableiten.
Es wird also ziemlich schnell deutlich, daß Protein- und DNS-Analysen auf ganz unterschiedlichen Niveaus ansetzen und somit eine unterschiedliche Aussagekraft haben. Die Bewertung der Ergebnisse muß also die benutzte Methode berücksichtigen.

Anwendungsbereiche der DNS-Analysen

Die hohe Aussagekraft von Sequenzanalysen soll im folgenden anhand einiger Beispiele dargestellt werden, die auch die große Breite der biologischen Anwendungsbereiche dieser Methode aufzeigen.
Wichtige allgemeinbiologische Erkenntnisse sind nicht nur für Verwandtschaftsanalysen auf Speziesniveau, sondern auch für Fragen der Großgliederung der Organsimen gewonnen worden. Die Schule von K. WOESE hat in jahrelanger Arbeit "ribosomale DNS"-Sequenzen bei Bakterien, Tieren und Pflanzen analysiert (WOESE, 1981). Bis dahin hat man die Bakterien als Organismen ohne Zellkern immer den sogenannten Eukaryonten (Tiere und Pflanzen) mit Zellkern gegenübergestellt. WOESE's Arbeit hat aber gezeigt, daß man bei den Bakterien zwei Gruppen unterscheiden muß, die Archaebakterien und die Eubakterien. Beide sind untereinander nicht näher verwandt als sie es zu den Eukaryonten sind (Abb. 3).

Andere Arbeitsgruppen haben DNS-Sequenzanalysen benutzt, um Verwandtschaftsbeziehungen innerhalb einer Gattung oder Familie aufzuzeigen. Als Beispiel sei hier eine Arbeit des Labors von A. WILSON vorgestellt, der als einer der Pioniere der molekularen Systematik gilt (MEYER et al., 1990). Der Viktoriasee in Ostafrika beherbergt 200 endemische Cichlidenformen. Die Frage, ob alle rezenten Arten von einer Stammart abstammen, oder ob es einen polyphyletischen Ursprung des Systems gibt, wurde lange Zeit diskutiert und konnte mittels mitochondrialer Sequenzanalysen beantwortet werden: Alle heute bekannten Formen sind monophyletischen Ursprungs.

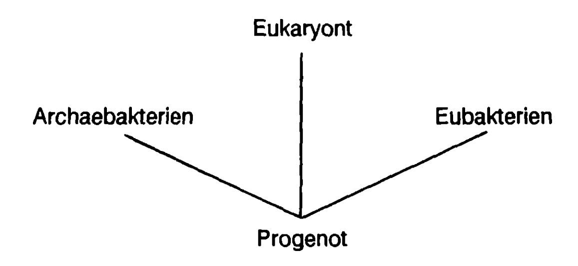


Abb. 3: Evolutionärer Stammbaum nach Untersuchungen der 16S und 16S-ähnlicher ribosomaler RNS bzw. DNS. Alle heutigen Lebewesen lassen sich in drei unabhängige Reiche eingliedern. Keines der Reiche stellt den Vorläufer für eine der anderen Linien dar. Heute geht man davon aus, daß alle drei Linien unabhängig von einem unbekannten Progenoten entstanden sind.

Diese kurzen Beispiele zeigen eindrucksvoll, daß DNS-Sequenzanalysen sehr präzise Antworten auf phylogenetische Fragestellungen liefern können. Es muß an dieser Stelle aber gesagt werden, daß der technische Aufwand, trotz zahlreicher wichtiger Verbesserungen und Vereinfachungen der letzten Jahre, im Vergleich zu den oben beschriebenen Proteinanalysen immer noch sehr hoch ist. Dennoch muß man aufgrund der großen Aussagekraft einer immer breiteren Anwendung der DNS-Technologie in der Systematik entgegensehen. Dies trifft sicherlich auch für die Lepidopterologie zu. In der Vergangenheit kamen molekulare Methoden nur selten und nur in Form der elektrophoretischen Techniken zum Einsatz. Nicht selten stehen die dabei erzielten Ergebnisse im Widerspruch zu den immer noch dominierenden morphologischen Arbeiten. Die Anwendung von DNS-Technologien sollte dazu beitragen können, bestehende Widersprüche zu klären und gleichzeitig neue Erkenntnisse auch in mikroevolutionären Fragestellungen zu erzielen. Die oben beschriebenen Beispiele zeigen eindrucksvoll, daß die Techniken dazu bereitstehen.

Morphologische und molekulare Evolution

Die zunehmende Anwendung molekularer Methoden in der Evolutionsforschung soll aber keineswegs morphologische Arbeiten verdrängen und als sinnlos erachten lassen. Vielmehr ist eine Synthese anzustreben, in der die verschiedenen Methoden – molekulare und morphologische Ansätze – parallel angewendet werden, um zu sicheren Aussagen über verwandtschaftliche Beziehungen zu kommen. Molekulare Evolutionsforschung hat nicht

nur in taxonomischen Fragestellungen eine Anwendung gefunden, sondern hat auch Einblicke in die Ursachen evolutionärer Veränderungen gewährt. Als Ausblick und Anregung soll am Ende ein Zitat des bereits erwähnten A. WILSON stehen:
"There may be two major types of molecular evolution. One is the process of protein evolution, which goes on at about the same rate in all species. The other is a process whoese rate is variable and which is responsible for evolutionary changes in anatomy and way of life. We propose that evolutionary change in regulatory systems accounts for evolution at and beyond the anatomical level" (WILSON et al., 1974:2843).

Literatur

Allgemeine Literatur zur molekularen Biologie und deren Methoden

ALBERTS, B. & L. JAENIKE (1986): Molekularbiologie der Zelle. – VCH Verlagsgesellschaft, Weinheim (2. Aufl. Garland, New York).

COOPER, G. (1980): Biochemische Arbeitsmethoden. – De Gruyter, Berlin.

HILLIS, D. M. & C. MORITZ (1990): Molecular Systematics. Sinauer Associates Inc., Sunderland, Massachusetts.

WESTERMEIER, R. (1990): Elektrophoresepraktikum. – VCH Verlagsgesellschaft, Weinheim.

WINNACKER, E.-L. (1986): Von Genen zu Klonen. – VCH Verlagsgesellschaft, Weinheim.

Spezielle Literatur

MEYER, A., KOCHER, T., BASASIBWAKI, P. & A. WILSON (1990): Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. – Nature **347**:550-553.

WILSON, A., MAXSON, L. R. & V. M. SARICH (1974): Two types of Molecular Evolution. Evidence from Studies of Interspecific Hybridization. – Proc. Nat. Acad. Sci. USA **71**:2843-2847.

WOESE, C. (1981): Archaebakterien – Zeugen aus der Urzeit des Lebens. – Spektrum der Wissenschaft, August 1981:75-91.

Anschrift des Verfassers:

Dipl.-Biol. RALF SOMMER
Institut für Genetik und Mikrobiologie
LMU München
Maria Ward Str. 1a
D-8000 München 19